Title: **Recovery of Trace Evidence in Forensic Archaeology**

**and the use of Alternate Light Sources (ALS)**

**Harte, A.a, Cassella, J. P.b, and McCullagh, N.A.c**

1. Department of Criminal Justice and Forensic Science, School of Law, Policing and Forensics, Staffordshire University, United Kingdom
2. Department of Criminal Justice and Forensic Science, School of Law, Policing and Forensics, Staffordshire University, United Kingdom
3. School of Natural and Built Environment, Queens University Belfast, United Kingdom

**Abstract**

The effectiveness of alternate light source (ALS) to fluoresce bone and other materials is well-attested to in a laboratory setting but rarely, if ever, has it been used in field excavation. This study examined the recovery rates of fragmentary bone, fabric, and metal, both with and without the use of an ALS, through practical and controlled excavation experiments with multiple users. All archaeology, including forensic archaeology and crime scene investigation more generally, should account for trace evidence. Currently, there is limited empirical data for the recovery of evidence from excavation, and those studies that do exist, highlight the short-comings in current methods.

Six comparable test pits were created, representing empty graves in which only trace evidence remained. Each contained 20 fragments of bone (≤10mm), 16 hair fibres, two pieces of fabric and two lead pieces, which were back-filled and left for over 15 weeks. Three excavators were each tasked with excavating two test pits: one using ALS, one in daylight conditions.

The results of the experiment identified some critical aspects of using blue 455nm wavelength ALS in the field, and the importance of experienced practitioners. Sample evidence was small in size and recovery rates were low. In daylight conditions, an average of 46% of trace evidence was identified, while just 40% was recovered using ALS. This excludes hair fibres which were almost undetectable in all conditions. When using ALS, smaller bone fragments were more than twice as likely to be recovered, but less non-fluorescent materials were found. The experience of each excavator had a positive correlation with excavation results. Excavation error rates were calculated, demonstrating that excavation is comparable using either technique, but daylight conditions lead to greater accuracy.

The findings suggest that ALS can be used to increase recovery of some evidence types. Test pits provided none of the usual primary evidence associated with graves and excavators had no prior experience of ALS. While retrieval rates were low, almost all recovered items were found *in situ* and an accurate records maintained. Error rates in forensic archaeology are essential and it is hoped that the method outlined here can be developed towards the establishment of acceptable error rates. While ALS use in forensic archaeology should not be considered a panacea to issues of trace evidence recovery, a combination of well-tested archaeological excavation methods, alongside the implementation of such proven forensic techniques, would likely lead to improved recovery of evidence.

**Keywords:** Forensic Archaeology, Alternate Light Source, Trace Evidence, Excavation, Bone, Fluorescence

1. **Introduction**

Recent studies have begun to highlight the necessity of gathering empirical data on the recovery of evidence in forensic archaeology (Evis 2016; Evis *et al,* 2016). Such research has indicated that the current methodologies and techniques are limited in their ability to recover all possible evidence. This has implications for the admissibility of excavated evidence within the legal process but should also be of concern throughout the field of archaeological practice. This limitation is particularly acute when it comes to smaller objects or trace evidence. The research outlined here examines the recovery of trace evidence from a controlled practical excavation experiment and assesses the application of an established technique in crime scene investigation – alternate light sources (ALS) – in a forensic archaeology context.

Forensic archaeology, the use of archaeological theory, methods and techniques in a legal context (Groen, Márquez-Grant & Janaway, 2015; NFI, 2013), has been developing as a profession since the 1970’s. Since the emergence of the discipline the focus of literature and practice has primarily been on the recovery of a body, or bodies, and seldom considers associated trace evidence to be of equal importance (Hochrein, 2011). At a crime scene, trace evidence is typically considered to include bodily fluids, hair, fibres, fingerprints, glass, paint, and explosive or gunshot residue (GSR) – simply, any material transferred during a crime (FBI, 2019) or items that provide information about the location, the victim or the suspect (Dotson, 2018). Consequently, fragmentary, or small disarticulated bone must also be considered as trace evidence. The significance of small bone has been highlighted during the investigation of mass graves or mass fatality incidents (Sturdy Colls, 2016; Fournet, 2016; Jugo & Wastell, 2015; MacKinnon & Mundroff 2007), mainly where there is evidence pointing to secondary and/or tertiary grave locations. The need to recover all remains from graves cannot be understated and such fragmentary remains have been used to verify pre-burial destruction of remains by explosives (Jugo & Wastell, 2015).

**1.1 Previous Studies**

Research on archaeological methods used in the forensic excavation of graves is largely a recent development. Tuller & Đurić (2006) retrospectively examined forensic excavations undertaken at two comparable mass graves in Serbia. The graves were excavated using two different methods, and quantified ‘unassociated’ bones recovered. The most suitable method was deemed to be the stratigraphic method, however, even in this case, small bones, from the extremities (*e.g*. phalanges, tarsals and carpels), were displaced frequently. These were more likely to be disarticulated, but more significantly, occasionally remained unrecognised and unrecovered by the excavators (Tuller & Đurić, 2006). There was seemingly no preservation issue that might explain this loss of smaller bones. As the methods were compared post-excavation, the experience levels of excavators, and the time they spent excavating, could not be meaningfully measured or compared.

Techniques used for excavating single clandestine graves have also been highlighted as a methodological concern. A recent study examined the outcomes of two different excavation methods in which eight experimental ‘graves’, each containing 20 artefacts, were excavated. The artefacts did not include any skeletal elements and were described as “items typically found in clandestine burials… also common, easily identifiable items, and… recognisable to participants” (Evis *et al*, 2016, 180). No further details of artefactual evidence or duration of burial were provided but the test pits were machine-dug (leaving toolmarks), were of ample size (1.2m x 0.75m x 0.85m) and contained 11 grave fills. Four archaeologists participated in the experiment, which resulted in stratigraphic excavation being identified as the most effective method, with an average retrieval rate of 71% (Evis *et al*, 2016.). The range of evidence recovery was between 55-82% (Evis *et al*, 2016*)*. There was also a positive correlation noted between the time spent excavating and the identification of items *in situ*. While archaeological experience ranged from one week to six years, the more experienced archaeologists recovered more collective evidence (Evis *et al*, 2016).

The most comprehensive examination into forensic excavation of single test graves has been that carried out by Evis (2016). Assessing four different excavation methods, recording systems and archaeological experience levels, the research consisted of 50 participants and as many test graves. Again, no skeletal elements were used, and evidence ranged from large items (*e.g*. dress, ID card) to small sized objects (*e.g*. earrings, hairpins *etc*.). The findings indicated that stratigraphic, demirant, and quadrant excavation methods all led to the recovery of over 70% of the evidential items. Of these, the quadrant method showed the least variability in results and was deemed the most reliable overall, yet small items were not always reliably identified, such as earrings (average 40% recovery) and hairpins (average 10% recovery), even when most participants opted to sieve the excavated soil. The total numbers of small items (*i.e.* earrings, hairpins, cigarette papers, coins and false nails) recovered *in situ* across these three methods was on average 41%. Two further aspects of Evis’s research are striking. The study found no positive correlation between the experience of the excavator and the successful recovery of items, except that non-archaeologists conversely found the least number of items *in situ*. Secondly, contrary to previous research, the time spent excavating a given test grave was not found to correlate with an increase in evidence recovered.

Undoubtedly these previous studies provide valuable insight and data into the success or otherwise of archaeological methods. The overarching aim of the research presented here, in line with previous research, is to “continue to experimentally test archaeological excavation methods… to ensure that they are suitable for use in forensic practice.” (Evis *et al*, 2016).

**1.2 Alternate Light Sources**

To date, the use of ALS has only been successfully applied in the search for scattered evidence on the surface. This has had the effect of increasing the recovery of skeletal remains by up to 70% (Brown & Christensen, 2018; Brown, 2017). This technique should be incorporated into methodologies previously proposed for the rapid initial assessment of surface human remains at a scene (Pokines, 2015; De Leeuwe & Groen, 2015). However, ALS as a technique has not previously been employed at excavations. Long-established field archaeology techniques have been adopted, uncritically and largely untested, as best practice in forensic excavation (Evis, 2016), while some recognised techniques of crime scene investigation have not been tested or applied at clandestine or mass grave excavations.

The use of alternative light sources (ALS) relies on photoluminescence, or fluorescence, where the object absorbs light of a lower wavelength and, through excitation, shortly after emits light of a greater wavelength (Miranda *et al*, 2014;Kersh *et al,* 2014; Viner *et al*, 2014; Christensen *et al*, 2014). Viewing the light emitted during fluorescence is controlled by the use of a long-pass filter, removing reflected incident light and competing light from background surfaces (Kersh *et al,* 2014). At crime scenes, this is well-established means of locating trace evidence and is both a simple and non-destructive technique (Miranda *et al*, 2014; Sheppard *et al*, 2017; Miranda *et al*, 2017; Lee & Khoo, 2010; Gallant, 2013; Schulz *et al*, 2007).

There is a high degree of overlap in the reported wavelengths that produce fluorescence in various materials, generally through blue and green spectral bands. While some materials such as hair, skin and fibres reportedly emit a fluorescent response at multiple wavelengths (‘Foster+Freeman’ 2019a; Viner *et al*, 2014; Schulz *et al,* 2007), blue wavelength ALS is optimal for the fluorescence of bone (Miranda *et al*, 2017; Swaraldahab & Christensen, 2016), GSR (Kersh *et al,* 2014; Atwater *et al,* 2006), blood (Foster+Freeman 2019a; Lee & Khoo, 2010) and other bodily fluids (Sheppard *et al*, 2017; Miranda *et al*, 2014; Zvarik *et al*, 2013 and Lincoln *et al*, 2006).

However, there is some variability in the excitation wavelength of bone. Human bone does fluoresce under UV-light, at 365 nm wavelength, (Bachman & Ellis, 1965; Hartles & Leaver, 1953), but blue light, *c*.450nm has proven to be more effective, even underwater (Christensen *et al*, 2014). It is accepted that the organic component of bone, collagen, accounts for the vast majority of the fluorescent response due to two amino acids,tyrosine and tryptophan (Swaraldahab & Christensen 2016; Christensen *et al*, 2014; Brown & Christensen, 2018; Bachman & Ellis, 1965). The collagen component forms a scaffold matrix for bone (Makareeva & Leikin, 2014), and is also necessary for radiocarbon dating and stable isotope analysis (Balzer *et al*, 1997). The inorganic hydroxyapatite also fluoresces under certain conditions, albeit weakly (Swaraldahab & Christensen 2016).

A recent study of ALS detection of human tooth and bone, found that that tooth material is more fluorescent than other bone and the optimal ALS wavelength for detection was at 455nm, viewed through an orange filter (Miranda *et al*, 2017). Notably, this analysis was performed against a background of inert Styrofoam, lacking the complexity of the environment experienced at forensic excavations or indeed other crime scene types.

A study carried out by Swaraldahab & Christensen (2016) concluded that bone fluorescence decreases over time. This study statistically assessed the performance of 490nm wavelength ALS with orange filter on bone of different post-mortem intervals – from recent (<5 years), to ‘ancient’ (mean = 522 years). Recent/semi-recent samples were from pig, while historic/ancient samples were human. There was an inverse correlation between the age of the sample and the intensity of fluorescence (*ibid*.). Although all bone of forensic interest fluoresced, the researchers cautioned that fluorescence is not a trustworthy indicator of ‘ancient’ bone (*ibid*.). This supports previous research where 2% of the forensically relevant samples were excluded as historic due to the absence of UV-fluorescence (Hoke *et al,* 2013).

The results of the study presented here evaluate the ability of archaeologists to recover small items of evidence from a buried context. As this sample evidence consisted primarily of bone fragments, but also hair fibres, cloth and metal, the application of ALS/fluorescence was also assessed. The controlled scenario was intended to reflect the reality of applying ALS in the field. Scavenging of skeletal material from shallow graves is also an issue that has been raised previously (Moraitis & Spiliopoulou, 2010). It was therefore important to record this activity, if it occurred, as such scavenging could have potentially impacted the quantification of results. The use of standardised test pits also allowed the accuracy of the excavations be examined, both in terms of samples and dimensions. The experience of the archaeologist, and the time spent excavating, were two other noteworthy factors to be re-evaluated in light of other studies (Evis 2016; Evis *et al.* 2016). The viewpoint of experienced practitioners in relation to the applicability of ALS in the field was also key to this project, particularly in terms of dark adaption and its effects on excavating, recording and photography. This report demonstrates the value of ALS in the identification of archaeological trace evidence and how future studies will further develop the technique to standardised procedures and methodologies for use in forensic archaeology.

**2 Materials & Methods**

**2.1 Materials**

Samples were chosen to reflect the potential of a ‘robbed-out’ primary clandestine grave, *i.e.* a burial that was reopened and the larger, anatomically positioned, skeletal evidence removed to another location (Boulestin & Duday 2006). This mostly consisted of fragmentary bone but also included hair fibres, fabric, and metal. No articulated skeletal remains were used (*i.e.* no individual skeletons) as is common in test pit experiments (Morse *et al*, 1976; Evis *et al*, 2016). All bone samples were of pig - *Sus scrofa domesticus –* a common substitute for human bone in research (Swaraldahab & Christensen 2016; Gallant 2013), and all bones selected were ribs. This fresh bone was defleshed, to increase the effects of soil directly on the bone and cut into fragments measuring 10mm3 and 5mm3.

Hair fibres were 50-70mm in length, undyed dark brown hair, and all sourced from the same individual, male, 30-40 years old. Cloth fabric was selected from a white garment of cotton/polyester mix. Each sample was created by cutting a square portion from the same garment, 2cm x 2cm. Lead was selected as the metal evidence as it could not fluoresce and therefore would act as a control for such potential evidence. Lead is also a common material encountered in ballistic evidence, and therefore significant within clandestine and mass graves (O’Brien, 2011). The lead material used was cut from roof-flashing, 10mm x 10mm and 3mm in thickness, and was square in shape – and so was not dissimilar to a distorted projectile. The size of these samples is notable in that all samples used, except for hair, are in the range 5mm – 20mm.

This sample evidence was buried within the test-pits for over 15 weeks (108 days), and control samples of each type were retained.

**2.2 Fieldwork Design & Procedure**

A couple of people that are standing in the grass

Description automatically generatedFor comparative purposes, the experiment design was largely informed by other recent forensic excavation recovery tests (Evis *et al,* 2016; Evis 2016). This needed to examine the rates of recovery of different types of trace evidence, both using stratigraphic excavation methods and using ALS. The first method was excavation in natural daylight, using single context recording. This also functioned as a control for analysing excavation using ALS. For comparative purposes, the set-up of each of the test pits was as near identical as was practicable. In all, six test pits were created, each from a template size – 0.6m x 0.3m in extent and 0.3m in depth. These were positioned adjacent to one another across the mid-slope of a south-facing ridge in agricultural grassland (Figure 1).

Figure 1: Layout of test pits, aerial view during excavation

Each test pit was rectangular, orientated north - south, and created using only a spade and hand shovel. The base of each of the 6 test pits was divided into 18 sectors (each approximately 100mm2) and an array of evidential material was laid out. This allowed the placing of sample evidence to be controlled, assuring that each test pit had identical items in similar positions. The sample evidence comprised 20 fragments of bone (10 x ≤10mm3 and 10 x ≤5mm3), 16 human hair fibers, two pieces of white polyester/cotton, and two pieces of lead. The array of sample evidence at the base of each test pit is shown in Figure 2.

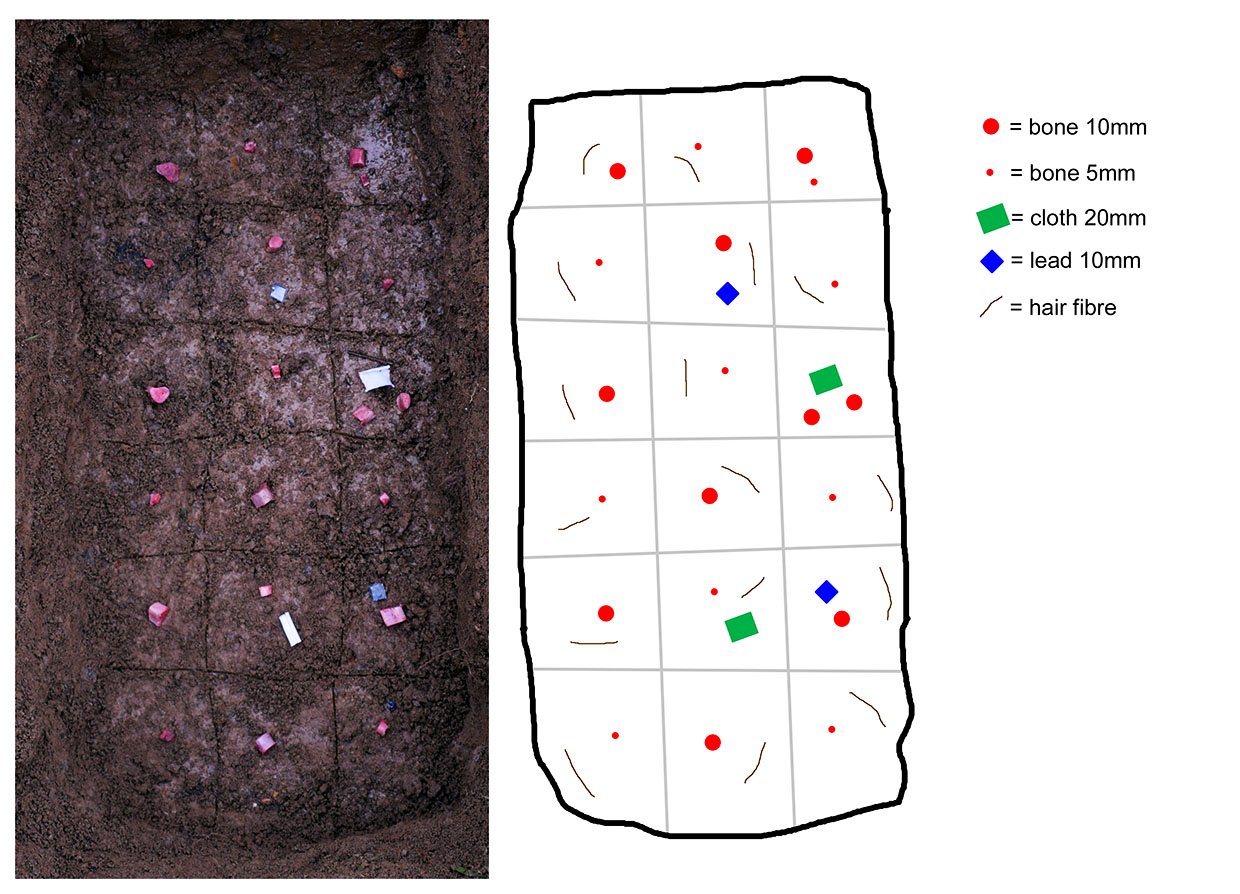
The test pits were backfilled with the same material that had been excavated out of each, a soft brown silty clay. This reflected the reality of any cut feature which might have been quickly infilled – as is the case in most clandestine graves. The original sod was replaced, to conceal the ‘graves’, and a wire mesh (2.5cm dia.) placed over the sod to deter potential scavenging, as samples did include fresh bone at a shallow depth. Scavenging was a factor that could impact the measurable outcomes of the experiment. Therefore, an infra-red motion-capture camera was also installed, focused on one test pit (No. 6), as an indicator of scavenger activity during the project. As the test pits had been created by the researcher, objective participants were required to excavate them. Three archaeologists of different specialisms and levels of experience volunteered, and their credentials are given in Table 1. Each participant was asked to excavate one test pit in natural daylight and another using ALS in a darkened environment (*i.e*. dark-tent). All excavation used the accepted stratigraphic/single-context method (Evis *et al.* 2016; MoLAS, 1994). Participants were not informed of the nature or position of sample evidence or the size, shape and fills of the test pits. They were each supplied with a brief to the project, outlining the tasks and the inter-confidentiality required between them during the experiment. This maintained objectivity and activities were alternated (*i.e*. participants A and C used ALS for their first test pits, while participant B used ALS for their second test pit).

Figure 2: sample evidence within Test Pit 5 (left); schematic evidence within sectors (right)

Table 1: Overview of participants

|  |  |  |  |
| --- | --- | --- | --- |
| Participant | Qualification(s) | Archaeological Experience (yrs.) | Profession |
| A | BA, MA, BMBS | 6 | Medical Doctor |
| B | BA, MA, MSc | 18 | Forensic Archaeologist |
| C | BA, MA, PhD | 21 | Osteo-archaeologist |

A record sheet for each test pit was completed contemporaneously with the excavation by each participant. These prompted the excavators to record standard archaeological information, including all evidence found.

Apart from standard excavation tools, a single ALS was the only scientific instrument in the field. This was a Sirchie Mini BluemaxxTM III LED. The technical specifications state this is a 3-Watt LED, emitting light at a wavelength of ‘approximately 455nm’ (Sirchie, 2017). Two pairs of Sirchie Orange filter goggles were also used. Photography of ALS on site used a Nikon D3300 with standard 18mm-55mm lens fitted with a Tiffen Orange long-pass filter. The dark tent was constructed of a timber frame with heavy-gauge grey polythene covering.

**2.3 Analysis**

The primary data analysis was quantitative, but observational information was also recorded. The quantitative analysis required a comparison of rates of evidence recovered between methods used (*i.e*. with ALS and without). Where this was marginal, or resulted in an ambiguous interpretation, an independent samples T-test was used to identify significant difference between methods. The overall numbers of recovered items were also analysed with respect to the experience of the excavator and the duration of the excavation procedure to see if trends emerged. Basic data analysis was performed in MS Excel (ver. 1905) with supplementary analysis (*i.e*. T-tests) conducted in IBM SPSS Statistics ver.25.

**3. Results**

**3.1 Quantitative Results**

The results of sample evidence recovery from each of the six test pits are as shown in Table 2. Note that these figures exclude hair fibres which were recovered in very low numbers, as discussed below. In no instance did hair fibres fluoresce in response to blue ALS.

Table 2: Recovered evidence from each test pit

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Test Pit (Excavator) | Lighting | Total items (exc. hair) | Total recovered (exc. hair) | Percentage (exc. hair) | In Situ | Ex situ | Mean % recovered |
| 1 (A) | ALS | 24 | 4 | 17% | 3 | 1 | 40.27% |
| 2 (C) | ALS | 24 | 14 | 58% | 11 | 3 |
| 3 (B) | ALS | 24 | 11 | 46% | 11 | 0 |
| 4 (C) | Nat Light | 24 | 14 | 58% | 12 | 2 | 45.83% |
| 5 (B) | Nat Light | 24 | 11 | 46% | 9 | 1 |
| 6 (A) | Nat Light | 24 | 8 | 33% | 8 | 0 |

The overall result was low, ranging from 17% - 58%. Excavation in natural light had a greater return on average, almost 14% more than when using ALS. A detailed breakdown of each type of sample recovered, and by which method, is given in Table 3. Notably, there was a similar overall result for larger bone ≤10mm and cloth fabric, while ALS provided a significantly greater return of smaller bone fragments. Other materials, namely those that did not fluoresce, were clearly more readily recovered in natural lighting conditions.

Table 3: Sample evidence types recovered from each trench

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Test Pit (Excavator) | Lighting | Bone ≤10mm | Bone ≤5mm | Cloth | Metal | Hair |
| 1 (A) | ALS | 3 | 0 | 1 | 0 | 0 |
| 2 (C) | ALS | 9 | 3 | 2 | 0 | 0 |
| 3 (B) | ALS | 6 | 2 | 2 | 1 | 0 |
|  | ALS Total | 18 | 5 | 5 | 1 | 0 |
| 4 (C) | Nat Light | 10 | 0 | 2 | 2 | 1 |
| 5 (B) | Nat Light | 4 | 2 | 2 | 2 | 1 |
| 6 (A) | Nat Light | 5 | 0 | 2 | 1 | 0 |
|  | Nat Light Total | 19 | 2 | 6 | 5 | 2 |

As there were an unequal quantities of sample types within each test pit, Figure 3 illustrates the mean percentages of each evidence type recovered by method. While marginally more bone in the region of ≤10mm size was found in natural light, more smaller bone fragments, ≤5mm in size, were found using ALS. Cloth was for the most part recovered but far less lead pieces (17%) were found when using ALS in comparison to naturally lit conditions (83%). Although all hair identified was found *ex situ*, both hair fibres (just 4% of the total) were found in natural light.

Figure 3: Percentage of each sample type recovered by method

An independent sample T-test was used to address the question of statistical difference between the two methods of excavation. The result is that Levene’s test indicates equality of variance (F= 1.196; sig.= 0.336) between the total samples recovered and that no statistically significant difference is noted between the two methods (T-test p =0.717). Even when size of bone is considered, no statistically significant difference was found between methods for either category of bone size (bone≤10mm sig.=.0.902; bone≤5mm sig.=.0.417). However, the identification of metal objects did show statistical difference (sig. = 0.047).

The excavated soil was not sieved at any point, and no item recovered could not be re-associated with the context, or area of the test-pit, from which it came. However, it was decided to record items that had been moved in the process of excavation as having been recovered *ex situ*. In both methods, 4 items were recovered in such a manner (Table 2), excluding 2 hair fibres, both found *ex situ* in daylight conditions. Therefore, *ex situ* evidence accounted for 14% of that recovered using ALS and 12.5% in daylight conditions.

In assessing the relationship between participants experience in archaeology and the level of sample evidence recovered a clear trend emerged. The more experienced the archaeologist the greater the recovery of evidence. However, in natural light the least experienced archaeologist returned more evidence while using ALS (Figure 4).

The time taken to excavate a given test pit was assessed in relation to the items recovered. The mean excavation times across both methods revealed that one item of evidence was recovered every 15 minutes. However, ALS seemingly impacted the rate of excavation – participant B was quicker while participant C was slower – significant as they both returned the exact same number of items in each method (see Table 4).

Figure 4: Experience (in years) and evidence recovery by method

**3.2 Observational Results**

During the 108 days that the test pits remained dormant, vegetation masked them completely. No scavenging was found to have occurred at the site during this project. The motion-capture camera did record foxes, magpies and cats, but there was no disturbance recorded (Figure 5).

A picture containing outdoor, grass, tree, street

Description automatically generated

Figure 5: Images from motion-capture camera with test-pit 6 at the centre. Although wildlife, such as foxes, were recorded in both day and night, no scavenging or other disturbance of the test-pits was recorded.

The practicalities of excavation meant that the sod covering each test pit had to be removed, and the pit cut identified at the top, prior to installing the dark-tent over it (as would likely be the case at any forensic scene). Note-taking within the darkened environment was also challenging, meaning that participants frequently breached the conditions required for optimum dark adaptation.

Participants noted that this ALS was cumbersome to hold while excavating, while a tripod in a set position cast shadows or was likely too distant/weak to examine the area being excavated. A higher-powered ALS unit may have overcome some of these issues.

The dimensions of each test pit, post-excavation, reflected the original measurements made by the researcher (see Table 4). This shows that experienced archaeologists can excavate such a cut feature in either daylight or a darkened environment with the same general accuracy. However, the rate of error is higher when using ALS due to the darkened conditions (error rate formula Original-Excavated/Original\*100). Notably, the least experienced participant (A) did not fully reach the base of the test pit in darkened conditions – with clear consequences to the recovery of sample evidence (*i.e*. 17%) but the calculable error rate does not reflect the true impact of this.

Table 4: Excavation time, original dimensions and excavated result of each test-pit, with calculated error rates. \* denotes which method was used first for each participant

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Method | Test Pit / Excavator | Time / min. | Original L x W x D (m) | Excavated L x W x D (m) | Actual Error (m) | Error Rate |
| ALS | 1 A\* | 125 | 0.56 x 0.28 x 0.30 | 0.56 x 0.25 x 0.26 | +0.07 | 6.14 |
| 2 C\* | 141 | 0.57 x 0.28 x 0.30 | 0.65 x 0.35 x 0.28 | -0.13 | 11.3 |
| 3 B | 177 | 0.56 x 0.28 x 0.30 | 0.63 x 0.32 x 0.30 | -0.11 | 9.65 |
| Average | 148 | Error with ALS | | ± 0.06 | 9.03 |
| Natural Light | 4 A | 125 | 0.58 x 0.26 x 0.30 | 0.65 x 0.35 x 0.28 | -0.14 | 12.28 |
| 5 B\* | 247 | 0.57 x 0.27 x 0.30 | 0.63 x 0.30 x 0.28 | -0.07 | 6.14 |
| 6 C | 125 | 0.56 x 0.28 x 0.30 | 0.56 x 0.28 x 0.30 | 0.00 | 0 |
| Average | 166 | Error in natural light | | ± 0.07 | 6.14 |

The excavation record sheets compiled by the participants showed that a satisfactory level of accuracy and recording was maintained in excavating the cut and fill in both methods. Only one mixed fill context was present, and this was accurately recorded by all excavators (in just one instance was soil colour not defined due to darkened conditions). The test pit cuts were all described accurately as rectangular with squared corners and flat bases.

Photography of *in situ* sample evidence as it fluoresced was challenging to achieve using ALS. It was largely restricted to close-ups only (Figure 6) and contextual photographs of the excavation were only achievable after completion once the tent structure was removed.



Figure 6: Excavation photography using ALS, close-ups of samples in Test Pits 2 and 3.

**4. Discussion**

In all findings, it must be highlighted that the archaeologists excavating the test pits did so without any prior knowledge of the size or shape of the cut or what items/objects the pits contained. Furthermore, there was no usual primary evidence (*i.e*. body or large objects) or investigative background, to inform the expected nature of the deposits. This meant that the archaeologists had no systematic template to aid their excavation strategies. However, the overall recovery of trace evidence was low, ranging from 17% to 58%. Even where the test pits were excavated in natural daylight, on average less than half of the available trace evidence was located (45.83%). The brighter conditions also promoted the diversity of evidence types recovered as not all items fluoresced. Hair fibres were only recovered in daylight, albeit *ex situ*, however, perhaps the best example was that of lead fragments, 83% of which were recovered, as against 17% when using ALS.

There was no statistical difference in overall recovery of objects between the two methods (p =0.717). It must also be reiterated that none of the archaeologists had prior experience of using ALS, or indeed of excavating in darkened conditions. Again, using ALS, less than half of the available trace evidence was recovered (40.27%). This is comparable to the rate of recovery of smaller items in a much larger study carried out by Evis (2016) but there is no directly comparable empirical data on the recovery of such trace evidence from test-pits or graves as the size of objects in similar experiments are generally larger. Notably, when using ALS, the participants were more than twice as likely to recover smaller bone fragments, *i.e*. ≤5mm. The importance of this should not be understated as it indicates the use of ALS would likely mitigate the magnitude of loss of small or fragmentary bones, as has been shown to occur in the excavation of mass graves (Tuller & Đurić, 2006).

The *ex situ* material recorded here was that which was moved during the excavation process but could still be associated with context. The majority of recovered material was found *in situ*, both using ALS and without. While the accuracy of the excavations was shown to be comparable between both techniques, the darkened environment while using ALS likely caused the incomplete excavation in one instance. This was the only instance of ‘under-cutting’ and so skews the average accuracy for excavation with ALS. Therefore, the error rate is more insightful, in that it was 47% greater when excavating in a darkened environment. There are no previously published error rates such as these and it is hoped that future forensic test excavations might replicate the method here, so that an acceptable error rate might be established. Overall, as might have been expected, excavating in darkened conditions using ALS is less accurate than in daylight.

The time spent excavating the test pits was generally between 2 and 3 hours. The exception was that of test pit 5, excavated in daylight by participant B, this took over 4 hours but did not result in a greater number of items recovered when compared to the same excavator using ALS, nor did represent the most items recovered overall. Similarly, time spent excavating did not indicate greater precision when defining the cut of the test pit. There is potential for bias due to familiarity while excavating the second test pit in each case. However, only participant A increased evidence recovered in their second test pit. Similarly, participants B and C shortened the time spent excavating their second test pits, and while excavation error increased for A and B, it improved for participant C (see Table 4).

Archaeological field experience appears to be an important factor. In assessing experience versus performance, Evis *et al* (2016) found the same positive correlation in the 1 to 6 years’ experience range. This study clearly demonstrates the same trend, as it continues from 6 years up to 21. It also shows that in daylight conditions the most accurate excavation was undertaken by the most experienced archaeologist and visa-versa. This contrasts with the larger study using 50 excavators in which no significant correlation was seen (Evis 2016), but it does seem that this assessed participants on their overall time spent in the profession, as opposed to their field experience only. Nevertheless, here as elsewhere, a positive correlation is recorded between experience and quantity of evidence recovered and therefore this significant aspect should not be ignored by the forensic and archaeology profession.

There are of course limitations in this experiment. Firstly, although six test pits were used, only outcomes from three archaeologists could be assessed. Secondly, as ALS has not been used in archaeological excavation previously, this compared an entirely unfamiliar technique alongside techniques that have been used for years by the participants. Finally, while the wavelength of the ALS instrument used was appropriate, the power output and width of the beam is likely to have been a limiting factor in excavations such as these. Darkened conditions and weak ALS also posed challenges in terms of the photographic record. Nevertheless, the duration that the samples were buried, and controls to record/prevent scavenging, meant that the test pits themselves were worthwhile, as both realistic and comparable.

Considering the results here, and elsewhere, the issue of trace evidence recovery in forensic archaeology remains largely unresolved. The results of this study indicate that ALS does increase the recovery of small and fragmentary bone. The depletion of collagen in bone over time is yet another factor to consider and is not only dependent on post-mortem interval (Swaraldahab & Christensen, 2016) but also soil type, sexual dimorphism, and age at death (Jellinghaus *et al*, 2019). This has potential implications for effective ALS use, particularly with regard to historic or cold cases. At this moment in time it appears to be impractical and irresponsible to excavate using only ALS, as the loss of non-fluorescing evidence types and the accuracy of the excavation would be unacceptable. However, in order to address the significant loss of evidence, ALS could be used on-site to identify small or fragmentary remains that would otherwise be discarded. This should inform excavation methodologies at a practical level. For instance, the excavated soil from each context could be sieved under ALS in a dark-tent nearby. Where smaller bone and other fluorescing items are shown to be over-looked by the excavator, the ALS/dark-tent could be moved over the burial periodically. Alternatively, once all visible remains have been removed the grave cut might be inspected with ALS as a precaution.

The more fragmented human remains are, the more essential it becomes to ensure that there is maximal retrieval of potential evidence and “that the remains are recovered by forensic archaeologists” (Scheuer & Black, 2007, 202). Therefore, a combination of well-tested excavation methods, such as stratigraphic, demirant or quadrant, alongside the implementation of established forensic techniques such as ALS in the field, would likely lead to better outcomes.

**5. Conclusion**

The current limitations in the recovery of trace evidence from buried environments needs to be further recognised and continue to be addressed in order to mitigate loss of material. This study has highlighted the use of ALS in archaeological excavation and provides a quantified baseline for potential errors in such excavations. No more than 58% of the sample items were identified by any participant, even in optimal conditions. This has potentially profound implications for the ability of archaeological excavation in a legal context to state the absence of evidence.

Whilst there are limitations in this experiment it must be considered alongside similar previous research. Even so, the successful recovery all available evidence is inadequate at best. Also significant in this experiment is the lack of previous experience in using ALS, a factor that may have caused some bias towards more traditional techniques. Nevertheless, excavation within a darkened environment using ALS did not statistically have an adverse impact on the amount of sample evidence returned. In fact, smaller bone fragments were identified more frequently, a distinct advantage of applying such a technique. It would likely be best to use ALS at different stages of the forensic excavation (as is done with metal detectors), particularly when nearing completion, to confirm that all trace bone or fabric is identified within the buried context. Alternatively, ALS could be used within an on-site darkened screening station, in which *ex situ* material could be recovered from excavated soil of a known context. Both methods would allow forensic archaeologists to further familiarise themselves with using ALS and perhaps lead to the refinement of the use of ALS in forensic archaeological excavations.

While the time spent excavating the test pits did not result in better outcomes, there was a clear correlation between the experience level of the excavator and amount of evidence recovered, whether in terms of sample items or the accuracy of the recorded feature. The error rates calculated for each test pit should decrease with larger features however the importance of establishing error rates in forensic practice cannot be understated.

Current methods and techniques of trace evidence retrieval in forensic archaeological excavation must be considered unsatisfactory, and the use of ALS, in a supplementary way, could increase overall recovery rates.

**Author contributions** (CRediT author statement)

**A. Harte:** Conceptualization, Methodology, Investigation, Resources, Analysis, Visualization, Writing - Original Draft. **J. P. Cassella**: Conceptualization, Supervision, Writing - Review & Editing **N.A. McCullagh**: Investigation, Writing - Review & Editing

**Declaration of competing interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The first author privately funded this work.

**Acknowledgements**: The authors thank Liz Deakin for related laboratory assistance, and the participating archaeologists for their commitment to the project.

**References**

Atwater, C.S., Durina, M.E., Durina, J.P., Blackledge, R.D. (2006) Visualization of gunshot residue patterns on dark clothing. *Journal of Forensic Sciences*, 51(5), 1091-1095.

Bachman, C., & Ellis, E. (1965) Fluorescence of bone. *Nature*, 206, 1328–31.

Balzer, A., Gleixner, G., Grupe, G., Schmit, H-L., Schamm, S., and Turban-Just, S. (1997) In Vitro Decomposition of Bone Collagen by Soil Bacteria: The Implications for Stable Isotope Analysis in Archaeometry. *Archaeometry,* 39(2), 415-429.

Boulestin, B., & Duday, H. (2006) Ethnology and archaeology of death: from the illusion of references to the use of a terminology. *Archaeologia Polona*, vol. 44, 149-169

Brown T.E. (2017) *Combining a Line Search with an Alternate Light Source when searching for Skeletal Remains*. Master of Science Research Project. George Mason University, Fairfax, VA.

Brown T.E. & Christensen A.M. (2018) Using an alternate light source (ALS) to search for surface deposited skeletal remains.  *Forensic Anthropology*, 1(1), 68-73.

Christensen, A.M., Horn, J.D. & Smith, V.A. (2014) The Use of an Alternate Light Source for Detecting Bones Underwater. *Journal of Forensic Sciences*, 59(4), 1046-1048.

De Leeuwe, R. & Groen, W.J.M. (2015) Forensic archaeology in the Netherlands: uncovering buried and scattered evidence. In Groen, W.J.M., Márquez-Grant, N., & Janaway, R. (eds.) (2015) *Forensic archaeology: a global perspective*. Wiley-Blackwell, London. 109-120.

Dotson, D. (2018) *List of Forensic Techniques*. [Online] October 2018. Available from: <https://sciencing.com/list-forensics-techniques-6392795.html> [Accessed: 12th May 2019].

Evis, L.H, Hanson, I., & Cheetham, P.N. 2016. An experimental study of two grave excavation methods: Arbitrary Level Excavation and Stratigraphic Excavation. *STAR:* *Science & Technology of Archaeological Research*, 2(2), 177-191. DOI: 10.1080/20548923.2016.1229916

Evis, L. 2016*. Forensic Archaeology: the application of comparative excavation methods and recording systems*. Archaeopress Publishing Ltd. Oxford.

FBI. (2019) *Trace Evidence* [Online]. Available from - <https://www.fbi.gov/services/laboratory/scientific-analysis/trace-evidence> [Accessed: 14th March 2019].

Foster+Freeman. (2019a) *Crime-Lite 82* [Online]. Available at <http://ffsupport.co.uk/Brochures/Crime-lite_82S.pdf> [Accessed 12th May 2019].

Fournet, C. (2016) (editorial) Forensic sciences (anthropology/archaeology/pathology) and international criminal justice. *Science and Justice,* 56, 303-304.

Gallant, A.S. (2013) Alternate Light Sources in the Detection of Bone After an Accelerated Fire: A Pilot Study. *Journal of Forensic Sciences*, 58(S1), 221-226.

Groen, W.J.M., Márquez-Grant, N. & Janaway, R. (eds.) (2015) *Forensic archaeology: a global perspective*. Wiley-Blackwell, London.

Hartles, R.L., & Leaver, A.G. (1953) The Fluorescence of Teeth under Ultraviolet Irradiation. *Biochemical Journal*, 54(4), 632-638.

Hochrein, M. J., (comp.) (2011) *A Bibliography Related to Crime Scene Interpretation with Emphasis in Forensic Geotaphonomic and Forensic Archaeological Techniques*. Fifteenth Edition. Published by author, Special Agent with the Federal Bureau of Investigation, Pittsburgh Division, Johnstown, Pennsylvania.

Hoke, N., Grigat, A., Grupe, G., & Harbeck, M. (2013) Reconsideration of bone postmortem interval estimation by UV-induced autofluorescence. *Forensic Science International*, 228(1–3), 176.e1–176.e6.

Jellinghaus, K., Urban, P.K., Hachmann, C., Bohnert, M., Hotz, G., Rosendahl, W., & Wittwer-Backofen, U. (2019) Collagen degradation as a possibility to determine the post-mortem interval (PMI) of human bones in a forensic context – A survey. *Legal Medicine*, 36, 96–102.

Jugo, A. & Wastell, S. (2015) Disassembling the pieces, reassembling the social: the Forensic and Political Lives of Secondary Mass Graves in Bosnia and Herzegovina. In Anstett, E. & Dreyfus., J-M. (eds.) *Human Remains and Identification: Mass violence, genocide, and the ‘forensic turn’*. Manchester University Press. 142-174.

Kersh, K.L., Childers, J.M, Justice, D., & Karim, G. (2014). Detection of Gunshot Residue on Dark-Colored Clothing Prior to Chemical Analysis. *Journal of Forensic Sciences*, 59(3), 753-762.

Lee, W.C. & Khoo, B.E. (2010). Forensic Light Sources for Detection of Biological Evidences in Crime Scene Investigation: A Review. Malaysian Journal of Forensic Sciences, 1, 17-27.

Lincoln, C.A., McBride, P.M., Turbett, G.R., Garbin, C.D., MacDonald, E.J. (2006). The use of an alternative light source to detect semen in clinical forensic medical practice. *Journal of Clinical Forensic Medicine*, 13, 215–218.

MacKinnon, G., & Mundroff, A.Z. (2007) The World Trade Centre – September 11, 2001. In Thompson, T., & Black, S., (eds.) *Forensic Human Identification: An Introduction*. BAHID. CRC Press, Taylor&Francis, London.

Makareeva, E., & Leikin, S. (2014) Chapter 7 - Collagen Structure, Folding and Function. In Shapiro, J.R., Byers, P.H., Glorieux, F.H. & Sponseller, P.D. (eds.) [*Osteogenesis Imperfecta*](https://www.sciencedirect.com/science/book/9780123971654)*: A Translational Approach to Brittle Bone Disease*. Elsevier Inc.

Miranda G.E, Haltenhoff-Melani R.F, Júnior L.F, & Júnior E.D. (2017) Use of an Alternate Light Source to Detect Tooth and Bone. *Brazilian Dental Journal*. 28 (1), 78-81.

Miranda, G.E., Prado, F.B., Delwing, F., & Júnior, E.D. (2014) Analysis of the fluorescence of body fluids on different surfaces and times. *Science and Justice*, 54, 427-431.

MoLAS, (1994) *Archaeological Site Manual*. Third Edition. Museum of London.

Moraitis, K. & Spiliopoulou, C. 2010. Forensic implications of carnivore scavenging on human remains recovered from outdoor locations in Greece. In *Journal of Forensic and Legal Medicine* 17, 298-303.

NFI, (2013) *Forensic Archaeology: uncovering buried and scattered evidence*. Netherlands Forensic Institute, Ministry of Security and Justice.

Pokines, J.T. (2015) A Procedure for Processing Outdoor Surface Forensic Scenes Yielding Skeletal Remains Among Leaf Litter. *Journal of Forensic Identification*, 65(2), 161-172.

Scheuer L. & Black S. (2007) Osteology. In Thompson T. & Black S. (eds.) *Forensic Human Identification: An Introduction*. BAHID. CRC Press, Taylor&Francis, London.

Schulz, M.M., Wehner, F. & Wehner, H-D. (2007) The Use of a Tunable Light Source (Mini-Crimescope MCS-400, SPEX Forensics) in Dissecting Microscopic Detection of Cryptic Epithelial Particles. *Journal of Forensic Sciences*, 52(4), 879 – 883.

Sheppard, K., Cassella, J.P., Fieldhouse, S. & King, R. (2017) The adaptation of a 360° camera utilising an alternate light source (ALS) for the detection of biological fluids at crime scenes. *Science and Justice*, 57, 239-249. <http://dx.doi.org/10.1016/j.scijus.2017.04.004>

Sirchie. (2017). *Technical Information: Mini Bluemaxx III Kit* [Online]. Available at - <https://www.sirchie.com/media/resourcecenter/item/b/m/bmk250a_ti08-595eng-rev2e_1.pdf>[Accessed: 16th March 2019].

Sturdy Colls, C. (2016) ‘Earth conceal not my blood’: forensic and archaeological approaches to locating the remains of Holocaust victims. In Dreyfus M-J. & Anstett, E. (eds.) *Human Remains in Society: Curation and Exhibition in the Aftermath of Genocide and Mass-violence.* Manchester University Press.

Swaraldahab M.A.H. & Christensen A.M. (2016) The Effect of Time on Bone Fluorescence: Implications for Using Alternate Light Sources to Search for Skeletal Remains. *Journal of Forensic Sciences*, 61(2), 442-444.

Tuller, H., & Đurić, M. (2006) Keeping the pieces together: Comparison of mass grave excavation methodology. *Forensic Science International*, 156, 192–200. doi:10.1016/j.forsciint.2004.12.033

Viner, T.C., Kagan, R.A. & Johnson, J.L. (2014). Using an alternate light source to detect electrically singed feathers and hair in a forensic setting. *Forensic Science International*, 234, e25–e29.

[Zvarik, M](https://www.ncbi.nlm.nih.gov/pubmed/?term=Zvarik%20M%5BAuthor%5D&cauthor=true&cauthor_uid=23790172)., [Martinicky, D](https://www.ncbi.nlm.nih.gov/pubmed/?term=Martinicky%20D%5BAuthor%5D&cauthor=true&cauthor_uid=23790172)., [Hunakova, L](https://www.ncbi.nlm.nih.gov/pubmed/?term=Hunakova%20L%5BAuthor%5D&cauthor=true&cauthor_uid=23790172)., [Lajdova, I](https://www.ncbi.nlm.nih.gov/pubmed/?term=Lajdova%20I%5BAuthor%5D&cauthor=true&cauthor_uid=23790172). & [Sikurova L](https://www.ncbi.nlm.nih.gov/pubmed/?term=Sikurova%20L%5BAuthor%5D&cauthor=true&cauthor_uid=23790172). (2013) Fluorescence characteristics of human urine from normal individuals and ovarian cancer patients. In [*Neoplasma*.](https://www.ncbi.nlm.nih.gov/pubmed/23790172) 2013;60(5):533-7. doi: 10.4149/neo\_2013\_069.